MTAR: A Robust 2D Shape Representation

Ibrahim El Rube', Systems Design Eng., University of Waterloo, Canada,
Naif Alajlan , Electrical and Computer Eng., University of Waterloo, Canada,
Mohamed Kamel, Electrical and Computer Eng., University of Waterloo, Canada,
Maher Ahmed, Physics and Computer Science Dep., Wilfrid Laurier University, Canada,
George Freeman, Electrical and Computer Eng., University of Waterloo, Canada

Abstract

In this paper, a new 2D shape Multiscale Triangle-Area Representation (MTAR) method is proposed. This representation utilizes a simple geometric principle, that is, the area of the triangles formed by the shape boundary points. The wavelet transform is used for smoothing and decomposing the shape boundaries into multiscale levels. At each scale level, a TAR image and the corresponding Maxima-Minima lines are obtained. The resulting MTAR is more robust to noise, less complex, and more selective than similar methods such as the curvature scale-space (CSS). Furthermore, the MTAR is invariant to the general affine transformations. The proposed MTAR is tested and compared to the CSS method using the MPEG-7 CE-shape-1 dataset. The results show that the proposed MTAR outperforms the CSS method for both tests.

I. Introduction

Shape representation is a crucial step in shape analysis and matching systems. The complexity and the performance of the subsequent steps in shape analysis systems are largely dependent on the invariance, robustness, stability, and uniqueness of the applied shape representation method. In the past decade, several methods were proposed for 2D shape representation and matching. Curvature scale space (CSS) [20],[1] and [18], fuzzy-based matching [8], dynamic programming [22], shape contexts [5], shock graphs [26], geodesic paths [14], Fourier descriptors [6], and wavelet descriptors [7] are examples of these techniques. A recent review paper can be found in [34].

The multi-scale approach for shape representation and matching is considered the most promising. It can be argued that human perception of shapes is a multi-scale by nature. In addition, many interesting shape properties are revealed at different scale levels. Another advantage includes its invariance to moderate amounts of deformations and noise. A popular and well established multiscale method is the Wavelet Transform (WT). It has been used in many disciplines including shape analysis and recognition [15], [9], [28]. Many researchers have adopted the Wavelet Transform (WT) in shape representation and matching. For example, WT was applied in 2-D domains in [16], [10], and [27], whereas WT was applied to 1-D shape boundary in [29], [30], [25], [11], [3], [13], [24], and [7].

Due to the spatial and frequency localization property of the wavelet basis functions, wavelet are more efficient in representing and describing shapes than Fourier descriptors and moments [7].

In this paper, a shape representation for closed boundaries that is called multi-scale triangle-area representation (MTAR) is proposed. The new representation enjoys many advantages over the curvature scale space (CSS) representation [20], which has been selected for MPEG-7 standardization after comprehensive comparative experiments with many other methods [21]. Unlike CSS representation where just the zero crossing points are invariant to affine transformations, MTAR is totally invariant to affine transformations. In addition, utilizing the wavelet transform makes the MTAR more robust to noise and provides selectivity in the matching process, that is, coarse-to-fine matching.

The remainder of this paper is organized as follows. Section II gives an overview of the related work. In Section III, the proposed representation is introduced. Section IV presents the matching approach. The experimental results are shown in Section V. Lastly, Section VI concludes our work and suggests future work.

II. Related Work

One of the most well-researched closed-contour shape representations is the curvature scale space (CSS) method proposed by Mokhtarian and Mackworth [19], [20], which has been selected for MPEG-7 standardization [21]. In their method, a Gaussian kernel with increasing standard deviation σ is used to gradually smooth the contour at different scale levels. At each scale, the curvature of each contour point is measured by:

$$c(u,\sigma) = \frac{\dot{x}(u,\sigma)\ddot{y}(u,\sigma) - \ddot{x}(u,\sigma)\dot{y}(u,\sigma)}{\left(\dot{x}(u,\sigma)^2 + \dot{y}(u,\sigma)^2\right)^{3/2}} \tag{1}$$

Where c is the curvature at location u and scale σ, $\dot{x}$ and $\ddot{x}$ are the first and second derivatives of x, respectively. By setting (1) to zero, the inflection points (or curvature zero crossings) are located at each scale. This results in a binary image, called CSS image, which shows the end-points of the concave segments along the contour (the horizontal axis) at each scale level (the vertical axis). As the scale level increases, the smoothing effect increases and the number of inflection points decreases until the contour becomes totally convex. Fig. 1 shows an example of the CSS image for a shape in the MPEG-7 dataset.

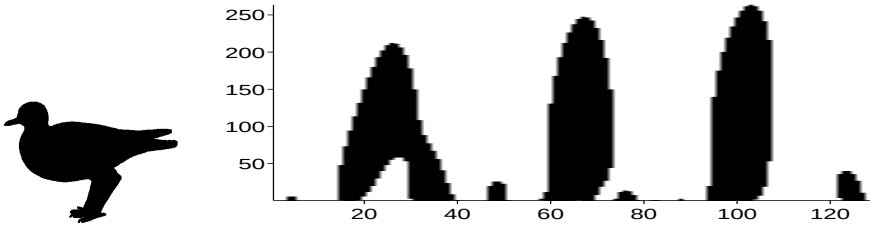

Fig. 1. An example of the CSS image for a bird shape.

Only the maxima of the CSS images' contours are used for matching two shapes [2]. Many heuristics are employed to find the best correspondence efficiently. This method exhibits certain degree of robustness to the affine transformation [1], where in another implementation explicit estimation of the affine parameters using the least squares approach was employed to verify the matching results [17]. The main limitations of the CSS method include its limited representation to only the concave segments and its failure to discriminate a shallow concavity from a deep one. For instance, totally convex shapes like squares and triangles have the same CSS image, which is only the background. Another disadvantage is that it requires a large number of scales to obtain the CSS image (may exceed 200).

Petrakis *et al.* proposed an approach for matching open and closed shapes using dynamic programming (DP) [22]. In their approach, implicit multi-scale matching takes place through matching merged contour segments in order to avoid the cost of computing the scale space explicitly. At first, the contours are approximated into polygons using cubic B-splines and inflection points are located, which segments the contour into convex and concave segments. Then, a set of geometrical features are computed for each segment to guide the merging process. The DP algorithm examines all possible merges of small segments of one shape to match with larger segments of the other and selects the best merge, i.e., that results in the minimum cost. An upper limit K is imposed on the maximum number of merged segments to compromise speed with accuracy. In case both shapes are closed, the complexity of the algorithm is $O(M^3N^2)$, where M and N are the number of inflection points of the two contours, and reduces to $O(K^2M^2N)$ when K is considered. The performance of this method was compared to that of the CSS method using about 1100 images of marine creatures. The CSS method achieved better accuracy only at small recalls (less than five retrieved shapes), otherwise this method is better. However, the dataset images are not labelled and the precision-recall assessment in this case is highly influenced by the human subjectivity. The main limitations of this method include the lack of robustness to the general affine transformations and the high

computational complexity of the matching process. In addition, polygonal approximation results in loss of some shape information and demands more computations.

An efficient computation of the curvature scale space representation using B-spline wavelets was proposed by Wang *et al.* [32], [31]. Their method provides an alternative to the classical Gaussian-based scale space representation while being much more efficient and relying on the well-established wavelet theory.

Another class of shape representation approaches is based on discrete Fourier transform (DFT) [33], [12]. They are mainly based on describing the contour by a limited number of coefficients in the frequency domain. A recent work of this category makes use of both amplitude and phase of the Fourier coefficients [4]. First, the low-frequency coefficients are normalized in terms of translation, scale and rotation. For matching, the inverse DFT is used to obtain normalized versions of the original contours in the spatial domain. Then, a DP method is employed to find the similarity distance between two transformed contours. Although this technique outperformed other Fourier-based methods, the authors reported less retrieval accuracy than the CSS method. Besides, it is not invariant to the general affine transformation, which makes it not suitable for 3D applications.

III. Multi-scale Triangle-Area Representation (MTAR)

In this section, the MTAR is introduced as a closed-boundary representation. MTAR has many desirable properties over other representations such as compactness, robustness to moderate noise and deformations, robustness to affine transformations, provision for efficient and effective matching, and handling partial occlusions. In addition, MTAR provides flexible coarse-to-fine matching. In the following, we give a brief explanation of how to obtain the TAR for an arbitrary closed contour. Then, the MTAR images using the WT is shown. Also the effect of general affine transformations on MTAR images is introduced.

A. TAR signatures

The TAR signature is computed from the area of the triangles formed by the points on the shape boundary. Each contour point is represented by its x and y coordinates. Then, the separated parameterized contour sequences $x(n)$ and $y(n)$ are obtained in order to facilitate applying 1D techniques on each sequence alone. The contour is re-sampled to N points and the curvature of each point is measured using the triangle area representation (TAR), as follows. For each three consecutive points P^1, P^2 and P^3, the signed area A of the triangle formed by these points is given by:

$$\begin{aligned} A &= \frac{1}{2}\begin{vmatrix} P_x^1 & P_y^1 & 1 \\ P_x^2 & P_y^2 & 1 \\ P_x^3 & P_y^3 & 1 \end{vmatrix} \\ &= \frac{1}{2}(-P_x^2P_y^1 + P_x^3P_y^1 + P_x^1P_y^2 - P_x^3P_y^2 - P_x^1P_y^3 + P_x^2P_y^3) \end{aligned} \tag{2}$$

When the contour is traversed in CCW direction, positive, negative and zero values of A mean convex, concave and straight-line points, respectively. Fig. 2 demonstrates these three types of the triangle areas.

For the complete boundary points, the TAR signature equation equals,

$$A(i) = \frac{1}{2}(-P_x^iP_y^{i-1} + P_x^{i+1}P_y^{i-1} + P_x^{i-1}P_y^i - P_x^{i+1}P_y^i - P_x^{i-1}P_y^{i+1} + P_x^iP_y^{i+1}), \quad i = 1:N \tag{3}$$

where N is the number of points on the shape boundary. The triangles at the edge points are formed by considering the periodicity of the closed boundary. Fig. 2 shows the complete TAR signature for the bird shape.

By Increasing the length of the triangle sides, i.e., considering farther points, the function of A will represent longer variations along the contour. Fig. 3 shows two examples of the TAR signatures plotted in 3D coordinates.

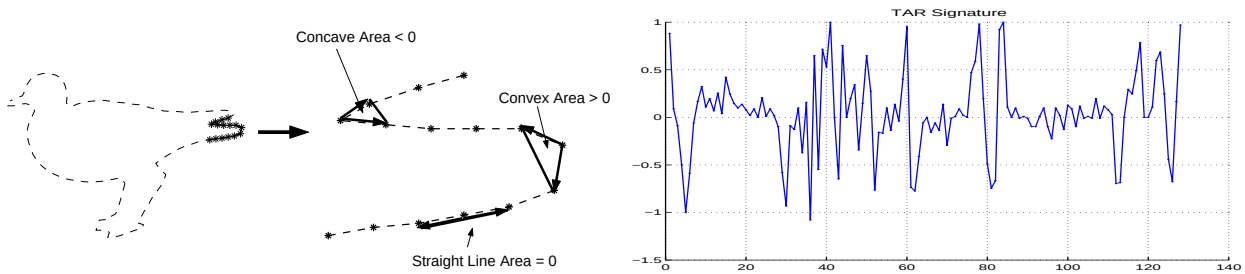

Fig. 2. Three different types of the triangle-area values and the TAR signature for the bird shape.

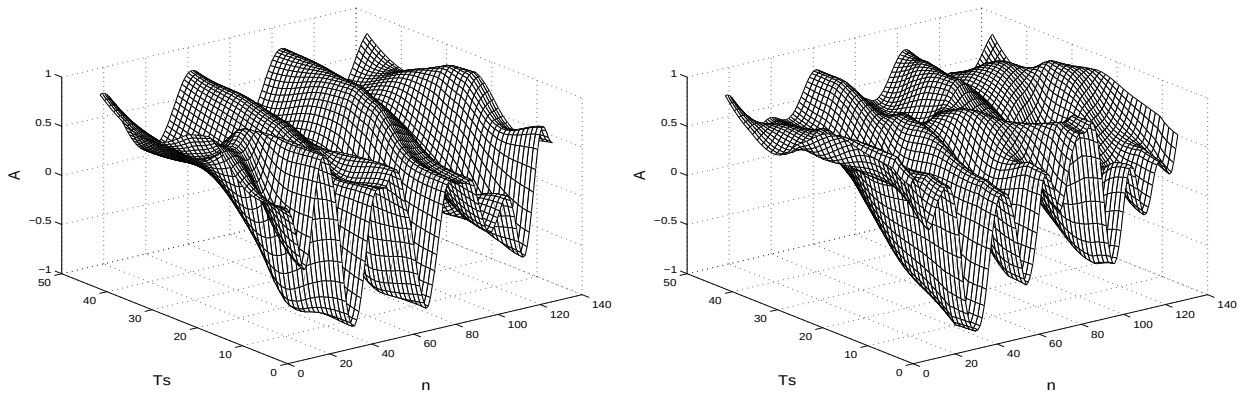

Fig. 3. 3D plot of the TAR signatures at the third scale for the bird and the Camel shapes, respectively. A is the amplitude of the TAR, Ts is the triangle-side length, and n is an index for the points on the shape boundary.

A TAR image is obtained by thresholding $A(i)$ at zero and taking the locations of the negative values at all values of triangle sides, as shown in Fig. 4. Thus, the horizontal axis in a TAR image shows the locations of the contour points and the vertical axis represents the triangle side length.

B. MTAR Images

The TAR image, like CSS image, is affected by noise. To reduce this effect, the dyadic wavelet transform is applied to each contour sequence obtaining various decomposed levels, as shown in the first column of Fig. 5. Only the approximation coefficients are adopted in order to reduce the noise effect. At each scale level, the same process of obtaining the TAR image in section III-A is also adopted here. The result is the MTAR, which contains $L + 1$ TAR images, where $L = log_2(N)$. Fig. 5 also shows that only small number of iterations (vertical axis) are required to obtain each TAR image. This is an advantage for the MTAR over the CSS method which required large number of iterations for its images.

The hierarchical representation of the MTAR is very useful when considering the coarse-to-fine matching, where fine-scale MTAR images can be used to eliminate dissimilar shapes and coarse-scale images provide discrimination between similar shapes. This property facilitates efficient search and enables MTAR to be applicable to search similar shapes in large databases.

C. Maxima and Minima points

Other important features that can be derived from the TAR signatures are the maxima and minima points of these functions. For simplicity, only the locations of these features are used in this paper at each triangle-side length. Fig. 6 illustrates a star shape represented by its concavity regions, maxima lines, and minima lines. The maxima points indicate the local convex points that correspond to a convex region. This

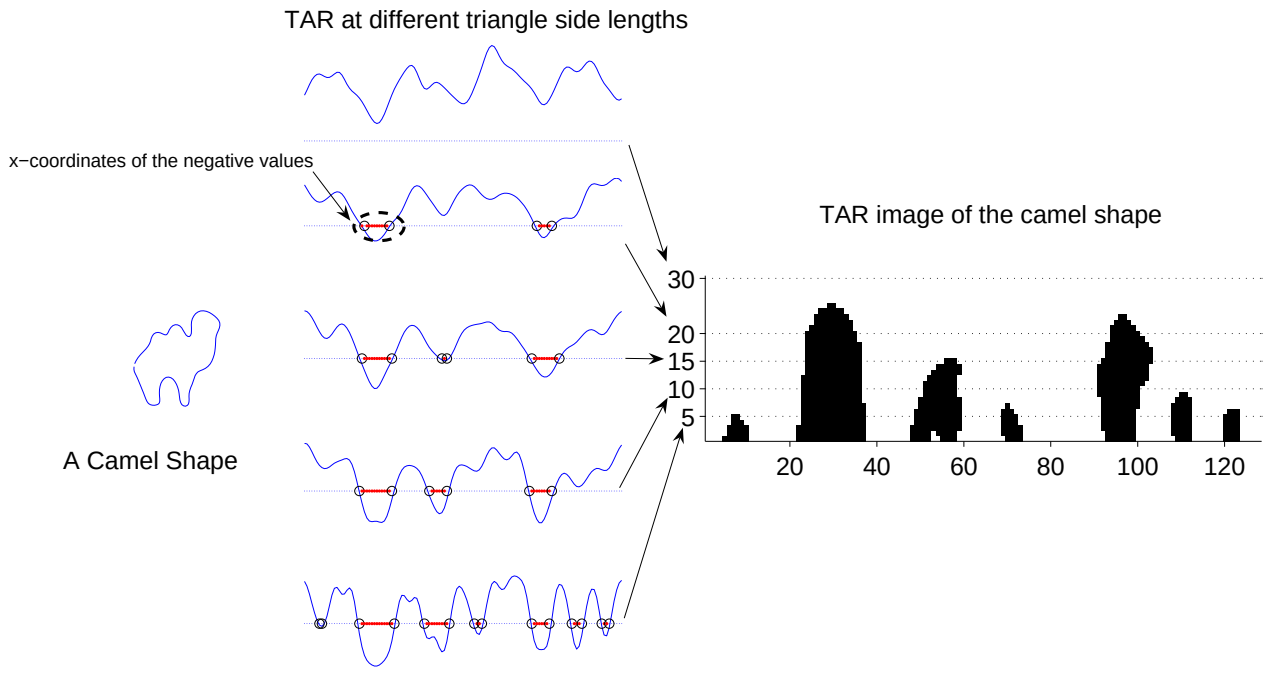

Fig. 4. Illustration of the computation of the TAR image. The second column shows the TAR signatures computed at specific triangle-side lengths. The TAR image (third column) is obtained using all the triangle-side lengths from 1 to 30.

region can occur in a totally convex area or within a concave region. For example, the small convexity region that is in the same direction as the East is located within a larger concavity. This is shown in the TAR image as a white (background color) region surrounded by the dotted $\bigcap$-shape concavity region. The minima points also could occur in both regions, as shown in the same figure. It should be noted that even when the shape become convex (no concavity regions), the local maxima and minima are still tracked. This is because that although the TAR becomes totally positive for large triangle-side lengths, it still has the minima and maxima segments.

D. Affine Transformations and MTAR

A desirable property of Eq.(2) is that it is invariant to affine transformations. Thus, MTAR images are also affine invariant. A general affine transformation applied to a shape is given by:

$$\begin{bmatrix} \hat{P}_x \\ \hat{P}_y \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} P_x \\ P_y \end{bmatrix} + \begin{bmatrix} e \\ f \end{bmatrix} \tag{4}$$

where $\hat{P}_x$ and $\hat{P}_y$ are the transformed coordinates of P_x and P_y, respectively, e and f represent translation and a, b, c and d reflect scale, rotation and shear. By substituting Eq.(4) into Eq.(2), we obtain:

$$\hat{A} = (ad - bc)\, A \tag{5}$$

where $\hat{A}$ is the affine transformed version of A. It is clear that $\hat{A}$, and hence TAR and MTAR, are relatively invariant to the affine transformations. Absolute invariance can be achieved by dividing $\hat{A}$ by its maximum value. Consequently, the values and the locations of the maxima and the minima points of the TAR signatures are invariant under general affine transformations (see Fig. 7).

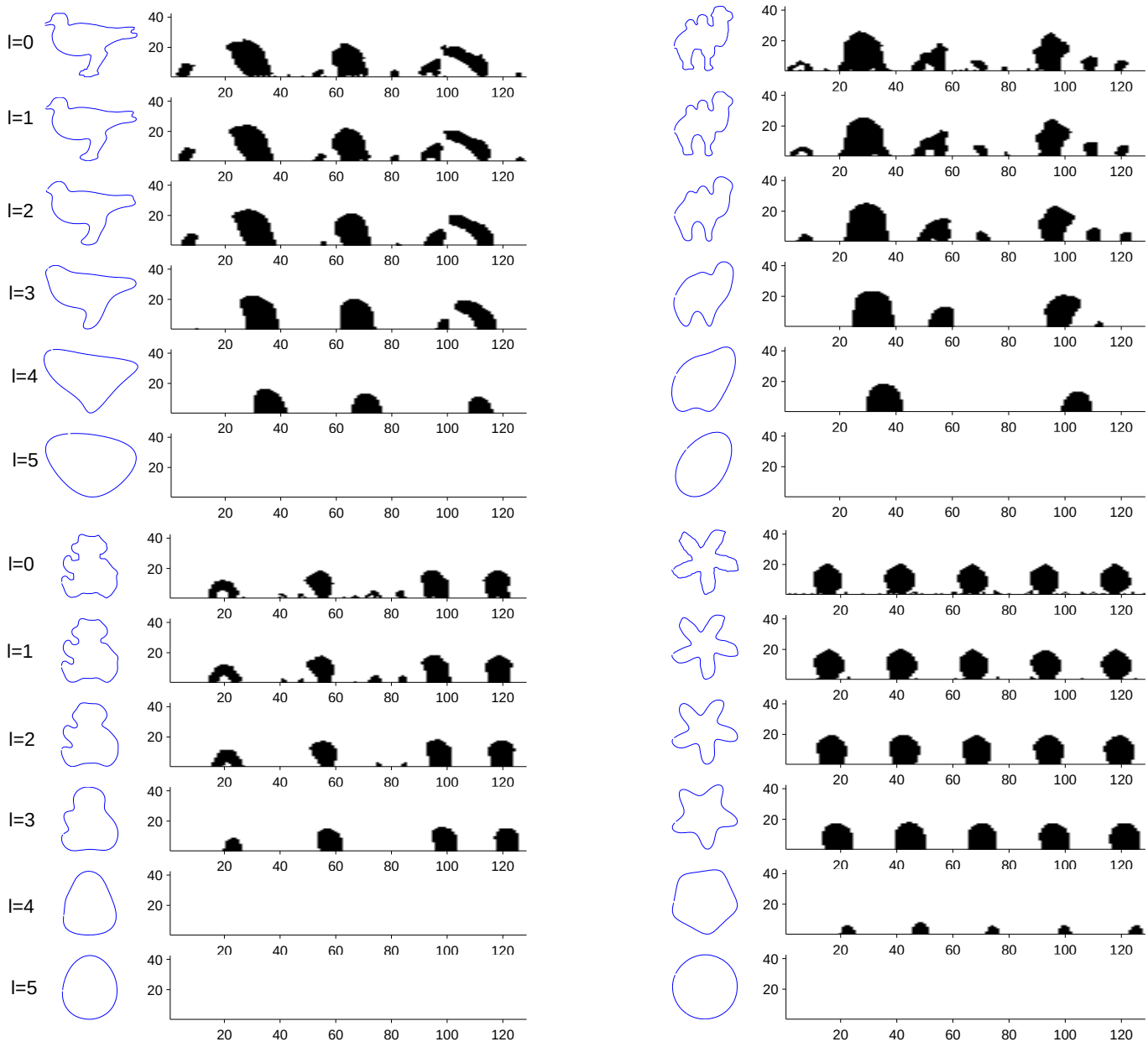

Fig. 5. Examples of MTAR images for different shapes. In each plot, the left column shows the approximated shape at different wavelet scale level (starting from l=0). The corresponding TAR images are shown in the right columns.

IV. Matching

There are some resemblances between an MTAR image and CSS image produced for the same shape. They both describe concavities along the contour. However, CSS method measures the curvature as the contour is smoothed by the Gaussian functions at different scales. On the contrary, each MTAR image represents the locations of the concavities using different triangle sides at a specific wavelet-smoothed scale level. Therefore, we followed a similar approach to that of CSS [23] in order to match two MTAR image sets of two shapes. Matching is performed in two stages as follows.

In the first stage, a set of global features are used to eliminate very dissimilar shapes and exclude them from further processing. In our work, these features include aspect ration AR, circularity C, eccentricity E and solidity S. In the second stage, a similarity measure D_s between each two MTAR images of the two shapes at certain scale is computed as described in [2]. D_s is based on finding a number of initial nodes or correspondences between two sets of maxima in the MTAR images using only two maxima in

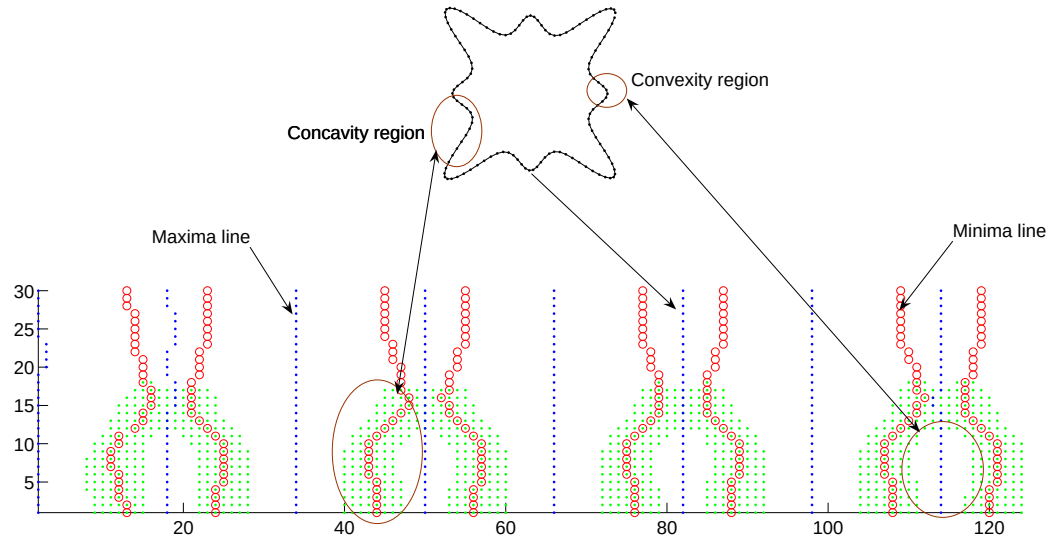

Fig. 6. A star shape and its TAR image showing the corresponding concavity region (dotted area), convexity region (background color), maxima line (dotted line), and minima line (circles line).

each image. Then, the lowest cost node is extended to include all other maxima and its cost is considered as D_s. However, we expand this approach to extend the two initial nodes with lowest costs and consider their minimum as D_s. This modification enhances the accuracy with a slight increase in computations.

As indicated before, the MTAR and the CSS images are concavity-based representation methods which mean that they can't differentiate between totally convex shapes (e.g., circles and squares) easily. In order to increase the discrimination of the MTAR method, the average sum of the maxima and minima points M_t for each TAR image is added.

The dissimilarity measure between each two MTAR images at a specific scale is a weighted sum of the global parameters, M_t, and D_s. Consequently, the final dissimilarity distance between two shapes is a weighted sum of the dissimilarity measures between their MTAR images. Practically, the MTAR images at scale zero (noise sensitive) and the last four scales (less informative) are avoided in calculations. Therefore, only four MTAR images are computed which reduced the computational complexity by more than a half. Note that MTAR images require less iterations to compute than CSS images.

V. Experimental Results

In order to test the proposed MTAR, the MPEG-7 CE-shape-1 dataset is chosen for the experiment. This dataset consists of 1400 shapes grouped in 70 classes. These shapes were taken from: natural objects, man-made objects, objects extracted from cartoons, and manually drawn objects. Recently, this dataset has been widely used for shape matching and retrieval.

Two experiments are conducted in this section. In the first experiment a retrieval technique based on the MTAR is tested and compared to the CSS-based retrieval algorithm. the second experiment investigates the invariance of the MTAR to the affine transformation distortions and compare it to the CSS method. To ensure a fair comparison between the two representations, the same matching algorithm is used.

A. Shape Retrieval test

In this test, the retrieval performances of the MTAR and the CSS methods are evaluated. The performances of the methods are assessed using the precision-recall curves, where a precision value at a

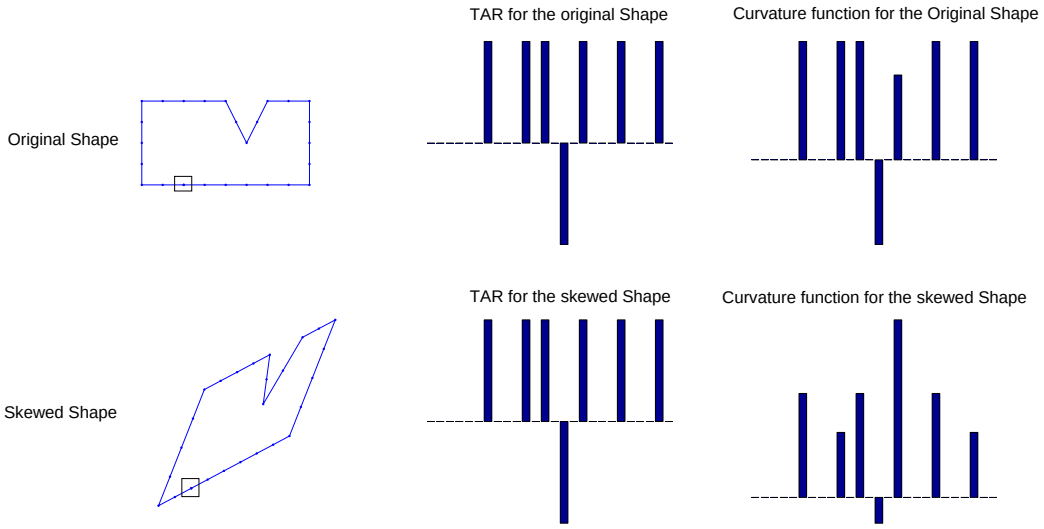

Fig. 7. A comparison between the TAR and the curvature function under affine transformation.

certain recall is the average of the precision values of all database shapes at that recall. Since the MTAR includes multiscale TAR images, the scale levels from $l = 1$ to $l = 4$ of the MTAR are evaluated and compared to each other. Fig. 8 show the precision-recall curves of these levels, where he performances of the first three scale levels are very close. These scales are affected by the degree of smoothing (i.e., the wavelet approximation level) which gives a trade-off between the robustness to noise and the degree of discrimination. Increasing the scale level (smoother shapes) increases the robustness of the shapes to the noise and to the small boundary deformations. However, the discrimination between shapes decreases, as all shapes tend to resemble to a circle or an ellipse at the end of smoothing. For instance, the shape at scale level $l = 4$ is more immune to noise than the previous levels, nevertheless it can barely distinguish between shapes.

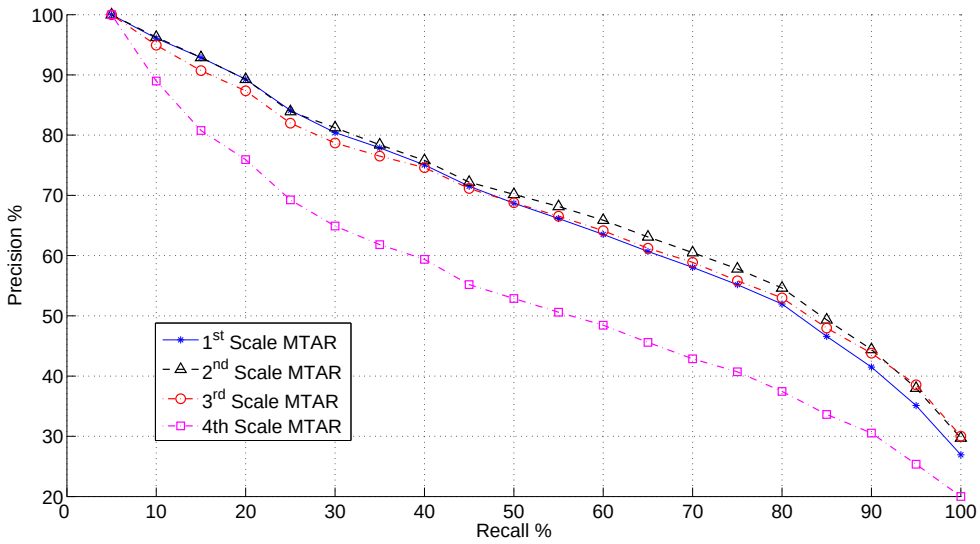

Fig. 8. Precision-recall curves for the scale levels $l = 1$ to $l = 4$ of the MTAR.

For the same four scale levels of MTAR ($l = 1..4$), the corresponding average sum of the maxima

and minima points is added and tested. Fig. 9 shows the plot of these scales. Although their plots look similar, it can be noticed from these two Fig.s that the first three scale levels of the MTAR have close precision-recall values. However, scale levels $l = 2$ and $l = 3$ are less sensitive to noise than the level $l = 1$ and hence needs less time for computation. Also the discrimination of these two scale is higher than the scale level $l = 4$. The performance of the MTAR is improved by using the average sum of the maxima and minima points, as seen in Fig. 9. This is due to the increase in the overall discrimination as a result of adding the convex feature points (i.e., the maxima and minima points). These points are important for distinguishing between totally convex shapes, since MTAR images as well as the CSS images are concavity-based representations.

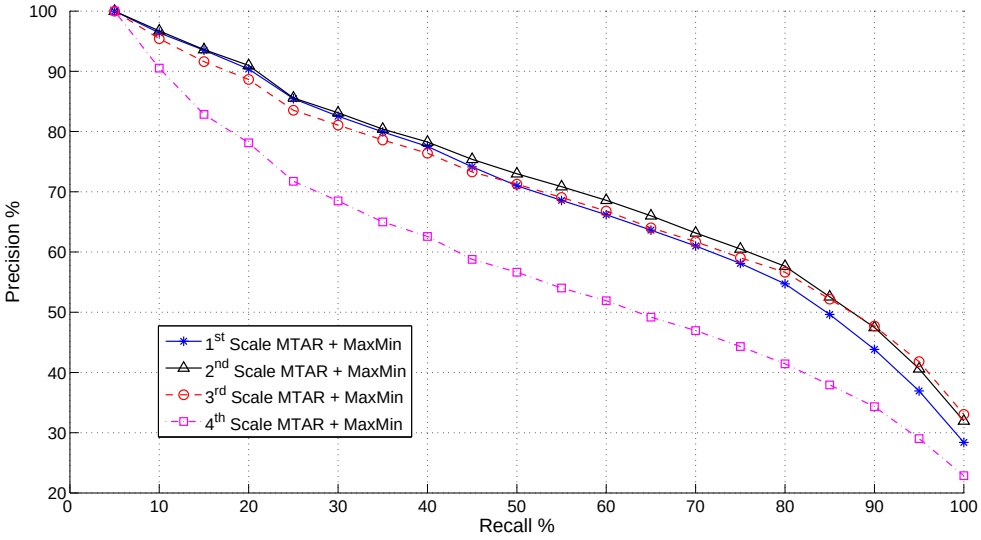

Fig. 9. Precision-recall curves for the scale levels $l = 1$ to $l = 4$ of the MTAR with the corresponding MaxMin.

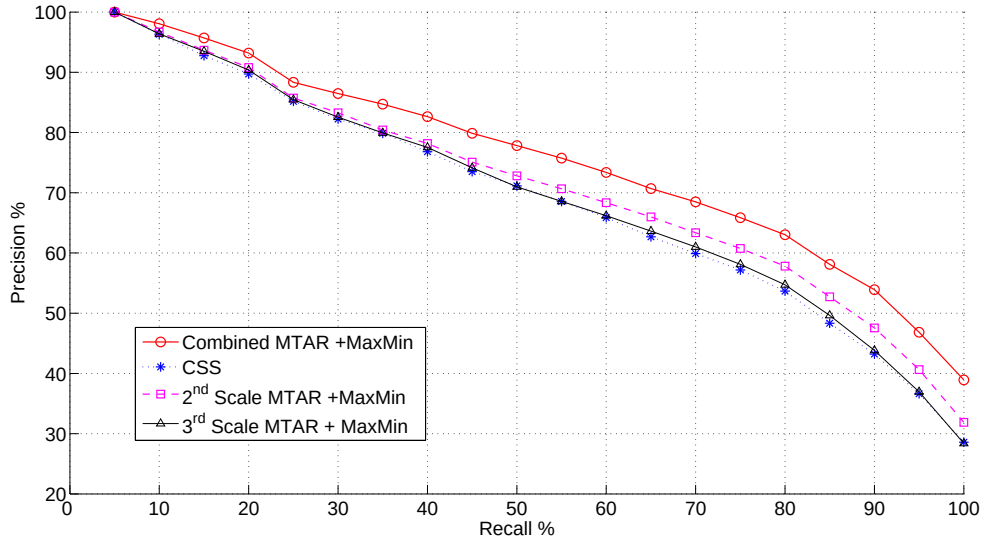

Fig. 10. Retrieval precision-recall curves for the combined MTAR+MaxMin, CSS, the 2nd scale of MTAR+MaxMin, and the 3rd scale of MTAR+MaxMin.

In order to assess the MTAR, it is compared with the similar method, the CSS. The CSS has been

selected for the MPEG-7 system after thorough and comprehensive tests for different shape descriptors. The results of Fig. 10 show that while the combined and the second scale have higher performance than the CSS, the third scale level has almost the same accuracy of the CSS with advantage of being less complex. This indicates that only one scale level can be used for shape retrieval with the advantages of less sensitivity to noise, and less complexity than the CSS.

The complexity of the MTAR method including the representation, feature extraction, and matching stages is tested by calculating the processing time required by each stage. Table I gives details about these times for both the MTAR and the CSS methods. The implementation of these methods is carried out using the Matlab©(ver. 6.5) program on Pentium IV 3.0 GHz PC. The listed representation and feature extraction times are the average times per single shape. For the matching stage, the times shown in the table are for the average time needed for a single comparison between two shapes.

Method	Representation	Feature Extraction	Matching
$1^{st} Scale$ MTAR	14.7ms	64ms	1.9ms
$2^{nd} Scale$ MTAR	15.6ms	35ms	1.5ms
$3^{rd} Scale$ MTAR	17ms	20.3ms	1.1ms
$4^{th} Scale$ MTAR	19.3ms	9ms	0.28ms
Combined MTAR	76.7ms	100ms	5.2ms
CSS	3382.7ms	174ms	1.8ms

TABLE I

PROCESSING TIMES PER SINGLE QUERY FOR THE MTAR AND THE CSS.

It is noticeable from this table that the MTAR is less complex than the CSS method, especially in the representation stage. The disadvantage of the CSS of being very time-consuming in obtaining the representation images does not appear in the MTAR method. This is because that the filtration in the CSS is applied at each iteration step until the all-convex shape is obtained. Also, for most shapes it requires more than 100 iterations to compute the CSS image. On the contrary, the MTAR is computed by directly applying equation 3 to the boundary sequences. Furthermore, each MTAR image represent only one scale level of wavelet smoothing. For example, to compute the 3^{rd} scale MTAR image, the wavelet smoothing filtration is applied only three times to the boundary sequences prior to the calculation of the TAR signatures at that level. The increase in the wavelet filtration iterations slightly increases the timing of the individual MTAR representations (in Table I) for higher scale levels. Conversely, the times for the MTAR feature extraction stages largely decrease for the corresponding wavelet scale levels. Also, the matching times decrease for the higher scale levels of the MTAR. This implies that the total average time required for a single comparison decreases for higher MTAR scale levels. However, since each scale level of the MTAR is computed independently, the complexity of the combined MTAR could be further reduced by using parallel-computing scheme.

B. Affine Invariance test

In this experiment, the invariance of the MTAR to the affine distortion is tested and compared to the CSS. The distorted shapes are obtained by transforming the original 2D images of 70 shapes representing the MPEG-7 dataset groups using Eq. (4). The parameters used to obtain these shapes are $b = [0, 0.4, 0.8, 1.5, 2, 3, 4, 6]$. A sample of these distorted shapes are shown in Fig. 11.

Fig. 11. Sample of the Affine Distorted Shapes.

Fig. 12 shows the precision-recall curves of the combined MTAR, the CSS, MTAR at $l = 2$, and MTAR at $l = 3$. The plots demonstrate that the MTAR achieves higher accuracy than the CSS, although the distortions in some shapes are high. Also, the single scale levels (i.e., at $l = 2$ and $l = 3$) have comparable performance to the CSS in the lower recalls whereas they exceed the CSS for higher recalls. This indicates that, in general, the MTAR is more robust to the general affine transformations than the CSS method.

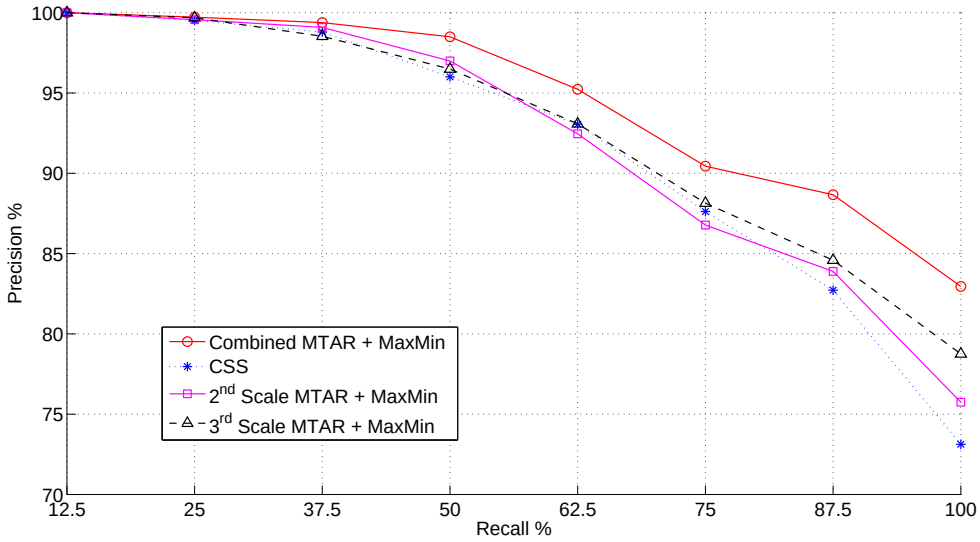

Fig. 12. Affine test precision-recall curves for the combined MTAR+MaxMin, CSS, the 2nd scale of MTAR+MaxMin, and the 3rd scale of MTAR+MaxMin.

VI. Discussions and Concluding Remarks

In this paper, a new multiscale shape representation method is introduced. The proposed representation utilizes the area of the triangles formed by each three consecutive and equally apart points on the shape boundary. The application of the wavelet transform in this method reduces the noise and the small boundary distortions, which accordingly improves the performance of the representation. The multiscale triangle-area representation (MTAR) images obtained from this method are used in shape matching and retrieval. The computed average sum of the maxima and minima points at each scale level increases the discrimination power of the proposed MTAR method. Two experiments are conducted in order to test the robustness and invariance of the proposed MTAR against the known image artifacts. The first experiment investigates the retrieval performance of the MTAR at various scale levels, while the second experiment tests the affine invariance of the MTAR method. When compared to the related representation method, the results show that, in general, MTAR outperforms the CSS method for both tests. Furthermore, using only one scale level (especially, when $l = 2$ or $l = 3$) attains comparable results and less overall complexity than the CSS method. The importance of comparing with the CSS is that it is already chosen as an MPEG-7 shape descriptor. Furthermore, several methods are compared to the CSS in the literature including the MPEG-7 experiment tests. In this paper, the MTAR outperforms the CSS even though using a matching algorithm similar to that in [23], which was originally designed for the CSS method. We believe that, the MTAR could achieve better performance by using a matching algorithm that takes into account more properties from its images (e.g., the directions and shapes of the concavities, the correlations between the maxima, and the correlations between the minima lines). Also, the performance of the MTAR is expected to achieve better performance when benefit from its hierarchical representation (i.e., coarse-to-fine matching), rather than using the weighted sum distances for computing the dissimilarity between shapes.

REFERENCES

[1] S. Abbasi and F. Mokhtarian. Affine-similar shape retrieval: application to multiview 3-d object recognition. *IEEE Trans. on Image Processing*, 10(1), January 2001.

[2] S. Abbasi, F. Mokhtarian, and J. Kittle. Curvature scale space image in shape similarity retrieval. *MultiMedia Systems*, 7(6):467–476, 1999.

[3] R. Alferez and Y. Wang. Geometric and illumination invariants for object recognition. *IEEE Trans. on PAMI*, 21(6):505–536, June 1999.

[4] I. Bartolini, P. Ciaccia, and M. Patella. Warp: Accurate retrieval of shapes using phase of fourier descriptors and time warping distance. *IEEE Trans. on PAMI*, 27(1):142–147, 2005.

[5] S. Belongie, J. Malik, and J. Puzicha. Shape matching and object recognition using shape contexts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(24):509–, 2002.

[6] T. Bui and G. Chen. Multiresolution moment-wavelet-fourier descriptor for 2-D pattern recognition. *Proceedings of SPIE-The International Society for Optical Engineering*, 3078:552–557, 1997.

[7] G. Chauang and C. Kuo. Wavelet descriptor of planar curves: Theory and applications. *IEEE Transaction on Image Processing*, 5(1):56–70, January 1996.

[8] Y. Chen and J. Wang. A region-based fuzzy feature matching approach to content-based image retrieval. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(9):1252–, 2002.

[9] L. Costa and R. Cesar Jr. *Shape Analysis and Classification, Theory and Practice*. CRC Press LLC, 2001.

[10] L. Feng and T.D. Bui. Classification of similar 2-d objects by wavelet-sparse-matrix (wsm) method. *International Journal of Pattern Recognition and Artificial Intelligence*, 15(2):329–345, 2001.

[11] R. Kashi, P. B-Kavde, R. Nowakowski, and T. Papathomas. 2-D shape representation and averaging using normalized wavelet descriptors. *Simulation*, 66(3):164–178, March 1996.

[12] H. Kauppinen, T. Seppänen, and M. Pietikäinen. An experimental comparison of autoregressive and fourier-based descriptors in 2d shape classification. *IEEE Trans. on PAMI*, 17(2):201–207, 1995.

[13] M. Khalil and M. Bayoumi. A dyadic wavelet affine invariant function for 2-D shape recognition. *IEEE Trans. on PAMI*, 23(10):1152–1164, October 2001.

[14] E. Klassen, A. Srivastava, W. Mio, and S. Joshi. Analysis of planar shapes using geodesic paths on shape spaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(3):372–, 2004.

[15] S. Mallat. *A Wavelet Tour of Signal Processing*. Academic Press, second edition, 1999.

[16] M. Mandal, S. Panchanathan, and T. Aboulnasr. Illumination invariant image indexing using moments and wavelets. *Journal of Electronic Imaging*, April 1998.

[17] F. Mokhtarian. Silhouette-based isolated object recognition through curvature scale space. *IEEE Trans. on PAMI*, 17(5):539–544, 1995.

[18] F. Mokhtarian and M. Bober. *Curvature Scale Space Representation: Theory, Applications, and MPEG-7 Standardization*. Kluwer Academic Publishers, 2003.

[19] F. Mokhtarian and A. Mackworth. Scale-based description and recognition of planar curves and two-dimensional shapes. *IEEE Trans. on PAMI*, 8(1):34–43, 1986.

[20] F. Mokhtarian and A. Mackworth. A theory of multi-scale, curvature-based shape representation for planar curves. *IEEE Trans. on PAMI*, 14(8):789–805, 1992.

[21] The MPEG Home Page. http://www.chiariglione.org/mpeg/index.htm.

[22] E.G.M. Petrakis, A. Diplaros, and E. Milios. Matching and retrieval of distorted and occluded shapes using dynamic programming. *IEEE Trans. on PAMI*, 24(11):1501–1516, November 2002.

[23] W. K. Pratt. *Digital Image Processing*. John Wiley and sons Inc, second edition, 1991.

[24] Ibrahim El Rube, Maher Ahmed, and Mohamed Kamel. Coarse-to-fine multiscale affine invariant shape matching and classification. In *ICPR (2)*, pages 163–166, 2004.

[25] Ibrahim El Rube, Mohamed Kamel, and Maher Ahmed. 2-d shape matching using asymmetric wavelet-based dissimilarity measure. In *ICIAR (1)*, pages 368–375, 2004.

[26] T. Sebastian, P. Klein, and B. Kimia. Computationally efficient wavelet affine invariant functions for shape recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(5):550–, 2004.

[27] D. Shen and H. Ip. Discriminative wavelet shape descriptors for recognition of 2-D patterns. *Pattern Recognition*, 32:151–165, 1999.

[28] M. Sonka, V. Hlavac, and R. Boyle. *Image Processing Analysis and Machine Vision*. PWS Publishing, 1999.

[29] Q. Tieng and W. Boles. Recognition of 2-D object contours using the wavelet transform zero-crossing representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(8):910–916, Aug. 1997.

[30] K. Tsang. Recognition of 2-D stanalone and occluded objects using wavelet transform. *International Journal of Pattern Recognition and Artificial Intelligence*, 15(4):691–705, 2001.

[31] Yu-Ping Wang and S. L. Lee. Scale-space derived from b-splines. *IEEE Trans. on PAMI*, 20(10):1040–1055, October 1998.

[32] Yu-Ping Wang, S. L. Lee, and K. Toraichi. Multiscale curvature-based shape representation using b-spline wavelets. *IEEE Trans. on Image Processing*, 8(11), November 1999.

[33] D. Zhang and G. Lu. Shape based image retrieval using generic fourier descriptors. *Signal Processing: Image Communication*, 17(10):825–848, 2002.

[34] D. Zhang and G. Lu. Review of shape representation and description techniques. *Pattern Recognition*, 37(1):1–19, 2004.